The Influence of Scale, Context and Spatial Preposition in Linguistic Topology

Anna-Katharina Lautenschütz[1], Clare Davies[2*], Martin Raubal[1], Angela Schwering[1,2] and Eric Pederson[3]

[1] Institute for Geoinformatics, University of Münster, Robert-Koch-Str. 26-28, 48149 Münster, Germany
annakl18@hotmail.com, raubal@uni-muenster.de, wiansc@wi.uni-muenster.de
[2] Ordnance Survey, Research & Innovation C530, Romsey Road, Southampton SO16 4GU, UK
clare.davies@ordnancesurvey.co.uk
[3] Linguistics Department, University of Oregon, Eugene, OR 97403-1290, USA
epederso@uoregon.edu

* (Author for correspondence)

Abstract. Following a similar method to that of Mark and Egenhofer (1994), a questionnaire-based experiment tested for possible effects of scale, context and spatial relation type on the acceptability of spatial prepositions. The results suggest that the previous assumption of scale invariance in spatial language is incorrect. The physical world as experienced by humans, and described by human language, is not a fractal: scale appears to change its very physical nature, and hence the meaning of its spatial relations. The experiment demonstrated how scale influences preposition use, and how different prepositions appeared to evoke different levels of acceptability in themselves. Context, in terms of object type (solid or liquid), interacted with these factors to demonstrate specific constraints upon spatial language use. The results are discussed in terms of figure-ground relations, as well as the role of human experience and the classification of the world into 'objects' in different ways at different scales. Since this was a preliminary and artificially-constrained experiment, the need for further research is emphasized.

1 Introduction

Spatial relations are considered to be one of the most distinctive aspects of spatial or geographical information. Despite occasional use of maps, diagrams and models, humans inevitably use language to communicate where objects are, and this is most commonly done by reference to their relation to other objects. Similarly, almost any GIS (geographic information system) query uses spatial relations to analyze or describe the constraints of spatial objects.

In order for GIS to be based on models of relevant geographic concepts, so as to improve usability and task relevance, in the early 1990s it was deemed important to develop formal models of the spatial relations that tend to exist within geographic space. Egenhofer and Franzosa [1] argued that spatial relations can be grouped into three different categories:

- Topological relations, which are invariant under topological transformations of the reference objects [2];
- Metric relations in terms of distances and directions [3];
- Relations concerning the partial and total order of spatial objects as described by prepositions such as ‘in front of’, ‘behind’, ‘above’ and ‘below’ [4].

Topological relations describe the spatial configurations of two objects, without reference to metric distance. For human spatial cognition, topology has long been considered the most important type of spatial relation, since Lynch [5] pointed out that humans remember urban topology and use it when wayfinding or navigating through space. According to Mark and Egenhofer [6], people capture and use topology more frequently and accurately than metric properties such as distance and shape. A recent analysis of the relations mentioned or implied in a national topographic dataset (Ordnance Survey of Great Britain's OS MasterMap® [7]) also showed that topological relations such as connection, intersection and adjacency were among the most commonly defined relations among geographic features. These relations are therefore the focus of this paper.

1.1 Spatial Language

How then are topological relations expressed in spatial language, such as through the use of prepositions? How predictable is the match between a given topological relation and people's choice of spatial preposition for describing it? This question is more important than it may at first appear. If we do not properly understand how spatial language is selected and interpreted by human speakers and listeners, then our GIS, robots and other technologies will not be able to reliably match the expectations and intentions of human listeners and speakers to linguistic spatial descriptions.

For example, in order to simplify the usage of GIS for non-experts it would be enormously helpful to be able to use natural language expressions; e.g. Riedemann [8] emphasized the importance of GIS terminology reflecting the user's language. Query languages can be improved when the predicates are chosen according to user needs, and the underlying cognitive understanding of spatial relations needs to be taken into account when defining user-appropriate semantics. In turn, this can improve the development of ontologies [9,10].

To progress towards this, we have to understand the factors that influence the choice of spatial terms (verbs and prepositions, and similar forms) when a speaker of a given language attempts to describe, or evaluate a description of, a specific spatial relation. Mark and Egenhofer [11] investigated this via an experiment with human participants. This examined the influence of geometric factors on people's acceptance of (agreement with) a sentence such as "The road crosses the park", when accompanied with drawings showing various configurations of the road, represented as a line feature, and a region (the park, represented as a featureless 2D region) that was partly or wholly 'crossed' by it. This study demonstrated that the notion of 'crossing' is not an all-or-nothing concept: rather there are degrees of acceptability of the tested linguistic term. For example, people would be less convinced of the sentence's validity where the line did not continue right across the region (e.g. if it doubled back).

More recently, a program of research by Kenny Coventry and colleagues (e.g. [12]) has demonstrated that the choice of preposition in describing spatial relations depends on other factors besides geometric configuration, at least at the immediate scale often referred to as 'tabletop' or 'figural' space. Two types of influential factor have been identified: 'functional' relations between two objects (such as a coffee pot and a cup), and 'dynamic-kinematic' relations (e.g. the apparently most likely movements that the objects will make relative to each other). This body of work suggests that geometry is not enough: the choice of spatial linguistic terms may depend on the nature of the objects under consideration.

One factor that has previously been ruled out of such considerations is that of spatial scale. Talmy [13,14] argued that when spatial language references topological relations, shape and magnitude are irrelevant to the appropriateness of the expression. Essentially, this manifests a claim that spatial language is scale invariant or *scale-neutral*: the same linguistic terms would describe the same spatial relations at any scale. This claim has been widely discussed by Talmy and others, e.g. at the NCGIA Specialist Meeting for Research Initiative 2 “Languages of Spatial Relations” where other attendees including Mark and Zubin questioned the results of scale neutrality [15].

It is important to know whether Talmy's claim of scale neutrality for spatial language is in fact a safe assumption, for the following reasons:

1. If the language used to describe topological spatial relations varies at different scales, this may suggest that people's underlying cognitive models of those relations may also differ in content, and hence in their availability for analogical reasoning and other aspects of problem solving [16].
2. Geographic information scientists need to know whether, and when, we can generalise from cognitive and linguistic studies of spatial language using figural spaces (e.g. items on a table) to environmental and geographical spaces [17]. If not, then research studies need to take scale explicitly into account before generalising about spatial language use.
3. This is also true for key findings such as those described by Coventry and Garrod - do non-geometric factors also come into play at larger scales? In particular, does the *context* of use, including the functional nature of the objects themselves, make a difference to the description of their spatial relations at some scales and not others?

We test Talmy's claim in the present study, and simultaneously examine the potential role of object type as just one, quite easily isolated, aspect of the context of use. Since we have no reason to assume that all spatial terms are equally affected by any given influential factor, we included three different topological relations of intersection, adjacency and connection. Following the same experimental paradigm as Mark and Egenhofer's seminal study, we focused on the relation between a line and an area (region) feature, as described in a sentence and illustrated by a simple diagrammatic drawing (which could be interpreted as a simple map at various scales). The rating of sentence 'acceptability' was the dependent variable. Talmy's claim forms a null hypothesis - that scale does not make a difference. Similar null hypotheses may be advanced for the effects of context and type of relation. In the next three sections we will explain how we distinguished among scales and object types, and how we chose which examples of spatial relations to test.

1.2 Scales in Spatial Perception

What do we mean by the concept of 'scales', and how can we differentiate among them? In general, a scale defines "the ratio between the dimensions of a representation and those of the thing it represents" ([17], p.313). Whereas this definition is meant for maps, scale in human perception is defined as "the size of a space relative to a person" (ibid).

Montello distinguishes between figural, vista, environmental and geographical space. *Figural* space is smaller than the human body and is apprehended without any locomotion, e.g. pictures, small objects and distant landmarks. *Vista* space, as the term implies, can be apprehended without locomotion, just by sight, but is larger than the human body. These spaces are usually single rooms, town squares and small valleys, but also include the surface of the earth as viewed from a plane. An *environmental* space (e.g. a city) is too big to be apprehended without locomotion and is learned over time; it does not need to be learned through models or maps, but these are often used as aids. *Geographical* space is much larger than the human body and is perceived over time mainly through symbolic representations, e.g. maps. Maps thus represent geographical and environmental space but are themselves a part of figural space, because they are much smaller than the human body.

The present experiment used Montello's scale distinctions, but whereas he distinguished four spaces, only three spaces are used here. The reason for this was the difficulty in representing scenes that were clearly and uniquely 'vista' space, as opposed to figural or environmental.

Montello's analysis suggested that human beings' perceptual experiences differentiate between these types of spatial scale. These different scales may therefore be reflected in people's choice of topological spatial language, if Talmy's claim of scale neutrality is false.

1.3 The Role of Context: Object Kinds or Geographic Feature Types

One of the most obvious aspects of the context surrounding a spatial relation is the nature of the two objects whose relationship is being described. Experience with geographic features and terms suggests that one of the most fundamental of these may be the distinction between solid and liquid features - in other words, between dry land and hydrology. The ways in which liquid objects behave, and hence their relations with each other and with solid objects, obviously differ from those among solid objects. In the geographic context in particular, but also at small scales where this distinction has generally been overlooked, this may be expected to make a difference to the way that spatial language terms are employed.

This distinction is also of interest to the domain of geospatial ontologies. Research into the potential of geospatial ontologies at Ordnance Survey and elsewhere is aimed at increasing the interoperability of geographical datasets, by adding semantics to enable comparison of concepts. Hydrology has been used as a major source of domain-specific concepts within this work to date [18], and it is important to know whether relations in this domain are likely to be distinctive in their manner of description.

2 Method

2.1 Experimental Design

As stated earlier, three factors were examined within this experiment – scale, context and type of topological relation – in an effort to begin to identify whether and when they influence spatial term use in natural language. The factor *scale* had three treatment levels: figural, environmental and geographic space. Figural space is the smallest scale used in the experiment: the examples chosen for this scale were string and trickle (line objects), for which the sentences described spatial relations to a leaf or a puddle (region objects). The next largest scale was environmental space, for which we used old road and stream (line objects), relating them to the region objects park and lake. For the largest scale, geographic space, we related the line objects gas pipeline and canal to the region objects country and sea.

The factor *context* (object type) has two levels, representing the two types of object – liquid and solid. The solid line objects at the three scales were string, old road and gas pipeline; the solid region objects were leaf, park and country. The corresponding liquid objects were trickle, stream and canal, and puddle, lake and sea. Each line feature at each scale was paired with the two different region features at the same scale.

In this initial experiment we wished to focus on spatial relations which were primarily topological, and of strong relevance to geographic information. Accordingly, spatial relations and prepositions were chosen that occurred frequently among the geographic features listed in the OS MasterMap™ Real-World Object Catalogue[1]. Prepositions that implied the third dimension or that were used in non-spatial contexts (e.g. 'on') were avoided, in order to focus clearly and simply on unambiguous spatial relations that could easily be represented both linguistically and diagrammatically. The *intersection* and *connection* relations were represented most often among such spatial prepositions in the catalogue, and were therefore chosen for analysis, along with the similarly common *adjacency*. It appeared from the catalogue evidence that these represented three of the most common exemplars of the various line-region relations described by the 'nine-intersection' model previously referenced by Mark and Egenhofer [6]. The three chosen relations also allowed us to keep a consistent sentence structure: all the relations were amenable to the same simple "The [line] runs [preposition] the [region]." This consistency of structure avoided introducing linguistic complications into the experiment, allowing a stronger and clearer test of the experimental factors of interest.

Each of these three relations was represented by two prepositions, to check whether different prepositions within a relation might yield different results. Again drawing on the spatial relations that we found to be most commonly used in the GIS context, the spatial relation *connection* was represented by the prepositions *from* and *to*; *intersection* was represented by the prepositions *across* and *through*; *adjacency* was represented by *next to* and *alongside*. The semantics of English prepositions are controversial and challenging to describe. Prepositions are either assumed to be highly polysemous [19] or semantically more general with complex rules of application [20]. However, for the purposes of the current design, every use of each preposition was designed to be as constant as possible, beyond the independent variable factors of scale and context. Accordingly, the results should be interpretable regardless of one's fuller semantic analysis for these English prepositions.

The participants were presented with 68 sentences describing a spatial relation. Each sentence was accompanied by a drawing to illustrate the spatial relation, in a similar style to those of Mark and Egenhofer [6]: see Figure 1. However, in the present experiment each of the three spatial relations was represented through only one drawing, used throughout the experiment. The drawings were intended to help to visualize the relation, and to demonstrate that despite the changes of scale and object, we were considering the objects more in terms of their spatial relations than their visual identities: this should have encouraged any tendency towards scale neutrality, and hence made a stronger test of our alternative hypothesis of scale dependence. It was explained that the drawings were not the key stimuli but merely illustrations, and participants were urged to focus on rating the sentence rather than the drawing itself[2].

[1] See http://www.ordnancesurvey.co.uk/oswebsite/products/osmastermap/

[2] Piloting the study showed that by making the drawings deliberately sketch-like, and choosing objects that could be fairly unconstrained in their shape and outline, the same drawing could serve as a reasonably convincing supporting diagram across all scales and objects, given that the instructions were to evaluate the sentence only.

For every combination of scale and spatial relation eight sentences were generated, to use all possible combinations of the line and region features and prepositions, leading to 3 x 3 x 8 = 72. However, this led to a problem with the spatial relation *intersection* because two liquid features cannot intersect, e.g. "the stream runs across the lake" would obviously be deemed unacceptable due to the physical properties of liquids. These 6 examples were therefore removed, while leaving the non-intersection cases of liquid to liquid relations, i.e.. adjacency and connection. This left 66 sentences in total. Two dummy (nonsense) sentences, deliberately mixing objects of different scales, were added in to test people's concentration (e.g. "The old road runs next to the leaf"): if participants rated these highly then their results were excluded from analysis. The stimuli were presented on paper; the experiment took about 20 minutes to complete. To check for potential order effects, participants were randomly assigned to complete the questionnaire either in its original form (within which the questions were randomly ordered), or with the question order reversed.

Table 1 illustrates how the various factors and examples were combined.

Table 1. Combinations of features in the sentences presented in the experiment: all line features were combined with all region features at their respective scales, using all six spatial prepositions apart from liquid-liquid intersections

	Line features	**Region features**
Figural	String/Trickle	Leaf/Puddle
Environmental	Old road/Stream	Park/Lake
Geographic	Gas pipeline/Canal	Country/Sea

The participants marked their agreement on a continuous line from 0 to 10: zero meant no agreement with the sentence; 10 meant that the sentence was perfectly accepted. Figure 1 shows two example sentences and drawings.

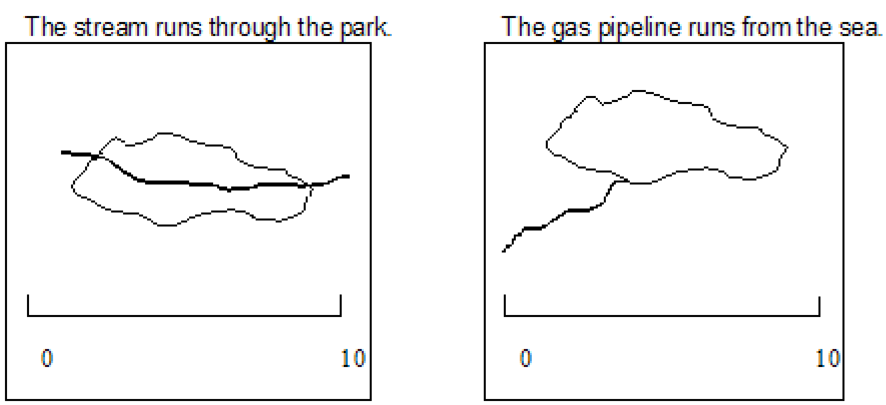

Fig. 1. Two examples from the questionnaire

2.2 Participants

26 adults completed the final version of the questionnaire. These were all volunteers living or working in the Southampton area of England. Three participants made unexpectedly high ratings of one or other of the two 'dummy' questions, a fourth participant was apparently only 10 years old, and two others did not have British English as their first language. These six were therefore excluded from analysis. This left 20 (an adequate sample in terms of statistical power for this entirely within-subjects design, given the expected effect sizes which were confirmed through piloting). The final sample included 9 females. Age ranged from 22 to 50 (mean=37 years, standard deviation=8). One female did not give her age. Otherwise the males were slightly but significantly older than the females (mean=40 as opposed to 32; t_{17}=2.89, p=0.01).

3 Results

Participants completed the experiment with no apparent problems or omissions, giving a mean rating of less than 1 to the two 'dummy' sentences but a mean rating of 7.0 (standard deviation=2.6) to the 66 test ones.

The results were analysed using a repeated-measures analysis of covariance (ANCOVA). Age was included as a continuous covariate, gender and question order as between-subject factors, and scale, preposition and object types as within-subject factors. None of the between-subject factors or the age covariate had any significant effects, so these will not be discussed further.

The main effects[3] of scale ($F_{2,1169}$=32.47, p<0.0001) and preposition ($F_{5,1169}$=7.41, p=0.0005) were both very strongly significant, as was the interaction between them ($F_{10,1169}$=3.52, p=0.0033). The main effect of object type was insignificant ($F_{3,1169}$=1.87, p=0.15). However, all of its interactions with the other variables were significant: scale x object type ($F_{6,1169}$=3.14, p=0.018), preposition x object type ($F_{13,1169}$=4.70, p=0.0021), and the second-order interaction scale x preposition x object type ($F_{26,1169}$=2.46, p=0.025).

The effect of scale showed that overall, across all object and relation types, sentences were rated higher at the environmental scale (mean=7.5, s.d.=2.4) than at the geographical (mean=7.1, s.d.=2.4) or figural (mean=5.9, s.d.=3.2) scales. The most highly rated preposition, across all scales and object types, was *through* (mean=7.5, s.d.=2.7), while the least was *from* (mean=6.6, s.d.=2.6) closely followed by *to* (mean=6.7, s.d.=2.7). As shown in Table 2, the interaction between the two factors was also particularly stark for *through*, which scored the lowest of all at figural scale, but the highest at geographical scale. However, this effect was not as strong for *across*, the other 'intersection' term. To a much lesser extent, *next to*, *to* and *from* also showed lower ratings at figural scale than at larger (especially environmental, as opposed to geographical) scales.

Table 2. Mean *(sd)* sentence ratings by scale and preposition

	figural	**environmental**	**geographic**
alongside	6.5 *(2.7)*	7.7 *(2.1)*	7.2 *(2.3)*
next to	6.0 *(3.0)*	7.2 *(2.5)*	6.8 *(2.4)*
to	6.2 *(2.8)*	7.2 *(2.5)*	6.4 *(2.6)*
from	6.1 *(2.6)*	7.2 *(2.3)*	6.3 *(2.6)*
across	7.0 *(2.5)*	7.2 *(2.7)*	7.3 *(2.3)*
through	6.0 *(3.2)*	8.0 *(2.4)*	8.5 *(1.4)*

The interaction of scale with object type suggested that at figural scale the sentences referencing a solid area (<u>leaf</u>) were rated much lower than at environmental scale (e.g. with solid-solid (<u>string</u>-<u>leaf</u>), mean=6.3, s.d.=2.8, but at environmental scale (<u>old road</u>-<u>park</u>), mean=8.1, s.d.=1.9). This was also true to a lesser extent for the other two object pairings. At environmental scale the solid-solid (<u>old road</u>-<u>park</u>: mean=8.1, s.d.=1.9) sentences were rated considerably higher than the liquid-liquid (<u>stream</u>-<u>lake</u>: mean=6.9, s.d.=2.8) and solid-liquid (<u>old road</u>-<u>lake</u>: mean=7.0, s.d.=2.8) pairings, with the liquid-solid (<u>stream</u>-<u>park</u>: mean=7.6, s.d.=2.2) sentences falling between the two. At geographical scale, however, all the object type pairings had similar average ratings of around 6.8-7.3.

The interaction of object type with preposition, across scales, suggested that the strongest-rated sentences were those using *across* with solid-solid object pairings (mean=8.2, s.d.=1.5), followed by those using *through* with liquid-solid (mean=7.7, s.d.=2.8). By contrast, the weakest sentences were those with a liquid-solid pairing and either *from* (mean=5.8, s.d.=2.7) or *to* (mean=6.0, s.d.=2.7), or with a liquid-liquid pairing and *next to* (mean=6.0, s.d.=3.0), or with solid-liquid and *across* (mean=6.3, s.d.=2.8). Thus *across* was shown to differ substantially between object types regardless of scale, whereas *through* differed substantially across scales regardless of object types.

[3] Conservative Greenhouse-Geisser significance values are quoted here, since sphericity was apparently violated for some main effects and interactions. However, this made no difference to the outcomes.

The significant second-order interaction (scale x object type x preposition) reflects the fact that certain individual sentences were particularly highly or poorly rated. Very low-rated sentences (i.e. averaging more than half a standard deviation below the overall mean, i.e. below 5.7) were as follows:

- *The trickle runs from the leaf* (figural, liq-sol, mean=4.9, s.d.=2.7)
- *The string runs through the leaf* (figural, sol-sol, mean=5.1, s.d.=3.3)
- *The trickle runs through the leaf* (figural, liq-sol, mean=5.2, s.d.=3.3)
- *The canal runs to the country* (geographical, liq-sol, mean=5.2, s.d.=2.5)
- *The trickle runs next to the puddle* (figural, liq-liq, mean=5.3, s.d.=3.6)
- *The old road runs across the lake* (environmental, sol-liq, mean=5.3, s.d.=3.4)

Very high-rated sentences (i.e. averaging more than half a standard deviation above the overall mean, i.e. above 8.3) were as follows:

- *The stream runs through the park* (environmental, liq-sol, mean=9.1, s.d.=0.7)
- *The old road runs through the park* (environmental, sol-sol, mean=8.9, s.d.=0.9)
- *The canal runs through the country* (geographical, liq-sol, mean=8.8, s.d.=0.9)
- *The old road runs across the park* (environmental, sol-sol, mean=8.5, s.d.=1.5)
- *The canal runs across the country* (geographical, liq-sol, mean=8.4, s.d.=1.4)

4 Discussion

The results of the ANOVA suggested that the participants did differentiate between scales in their response to specific prepositions. The overall findings show that more sentences were rated inappropriate in *figural* space than in *environmental* or *geographical* space. Therefore, the null hypotheses of scale invariance in use of spatial relational terms can be rejected. Prepositions which are accepted at the scale of local geographic space do not scale down easily to tabletop-scale items. Similarly, we can reject the assumption that connection, adjacency and intersection terms have similar acceptability across scales and contexts: *from* seems to be more problematic overall than *through*. Yet the ratings of *through* varied with scale more than any other preposition we tested, while *across* did not vary across scale but did vary across object type.

A moment's reflection by the reader will confirm that in fact *across* and *through* indicate very different relations when used for manipulable objects, and are only frequently interchanged in English at the environmental or geographical scale. Since *across* implies that the ground is a surface, while *through* tends to more strongly imply the third dimension, acceptability ratings can be expected to vary more for the latter in relation to changes in the perceived shape/substance of the 'ground' object (i.e. that through which the figure object 'runs'). Later we will further discuss the relevance of such figure-ground distinctions.

Environmental space was the best-rated scale overall: most sentences were rated positively, and two out of the three best rated sentences belong to this scale. By looking at the three best and three worst rated sentences in this scale, it becomes obvious that *context* has a major effect on the participant’s rating. All of the three best rated sentences have a solid region feature, namely park, and all of the three worst rated sentences have a liquid region feature, namely lake.

What factors might cause these effects? First, regarding the differences among prepositions, it appears that the 'connection' terms *to* and *from* were least transferable across object types. This is probably because people expect edge-to-end connections to be specified only where the two connecting objects are of the same type. For instance, a stream may run from/to the edge of a lake (or a canal from/to a sea), but for it to run 'from' a park one would assume that it began somewhere within the park, not at its border; similarly with canal and country, and with trickle and leaf.

To and *from* may also be problematic with solid-solid connections where the objects have no obvious causal relationship, e.g. the string running to/from the leaf, or the gas pipeline appearing to start/end abruptly at the border of the country. Yet the old road running from or to the park - perhaps because one might easily imagine a road that ends at the park's edge - is less obviously problematic. The sentences are rated negatively if a canal or gas pipeline runs *to* or *from* a country. Sentences are not rated that negatively when the canals or gas pipelines run *to* or from the sea. Here people's reasoning may be shaped through their experience with maps and models. Images of world maps or country maps show very often how a line feature ends or starts at the sea (or appears to do so, e.g. if it subsequently runs underwater). Conversely, they do not tend to end or start at a country. This suggests that *to* and *from* would only have been deemed acceptable if the region object was a physically different kind of thing from its surroundings - i.e. sea - which is apparently not a constraint that people expected at the small scale. So again, scale made a difference to the expected physical circumstances implied by a given preposition.

This links to the most difficult issue we encountered when designing and analysing this study. It is hard to assess the effect of scale on preposition use, independently of object type, because *objects at different scales have different characteristics*. At figural scale, it proved extremely hard to find an irregularly shaped 'area' object that might be deemed topologically equivalent to country or park - i.e. something relatively two-dimensional, that could be traversed over and beyond within the same plane. The smaller the space, the more items tend to be understood as individual, non-continuous objects with a salient third dimension. By contrast, in larger spaces solid areas tend to be easier to consider as a flat surface, not least because they tend to be considerably wider than they are tall. Thus *through* was considered unacceptable with leaf (implying penetration or depth perpendicular to the plane of the leaf), but unproblematic both with park and country (implying only traversal). In addition, larger-scale objects tend to be more continuous with others, and their boundaries increasingly abstractly rather than physically defined; thus the apparent unlikelihood of a gas pipeline stopping abruptly at a country's border.

The results concerning object type pairs have important implications for the use of spatial prepositions in describing object relations between hydrology and other domains. The characterisation of these relations is a key factor in ongoing work at Ordnance Survey and at Münster, to formally model hydrological and topographic data ontologies (e.g. [18]) The apparent lack of overall differences between object type pairs seems to have been due to different effects (at different scales and relation types) cancelling each other out, rather than to an indifference of spatial semantics to the physical state of objects. The same spatial relationship is not implied by the same term, when the objects are water rather than solid features. A liquid connecting to a solid is not deemed realistic if it is depicted as connecting with the edge, rather than showing a termination point inside the solid area. The issue here is people's expectation of *physical realism*.

For *across*, a solid line object crossing a solid area was seen as plausible at all scales, but not when it crossed a body of water (puddle, lake or sea). Again, this relates to expectations of physical realism: the old road or gas pipeline (but perhaps not the string, which was less harshly rated) would require some form of suspension over the water. Since the form of this suspension (e.g. a bridge) was not explicitly drawn or stated, participants found the relationship less plausible. Similarly, when a liquid line feature (trickle, stream or canal) was declared to be *next to* or *alongside* a water body (puddle, lake or sea), participants may have expected to either see a clear spatial separation between the two, or to have one mentioned in the sentence.

This has implications for cartographic representation, as well as for choice of linguistic terms when describing adjacent hydrological features. Physical structures that must be in place to avoid two water features merging together, or a solid line feature collapsing into a water body, need to be specified explicitly if the relevant spatial relations are to be deemed plausible.

Another issue concerning scale is the way in which different spaces tend to be apprehended. Geographic space is not only apprehended through physical actions such as manipulation and locomotion, but also with the help of models and maps. In fact, at both environmental and geographic scales, the images we used would have been interpreted as small-scale maps rather than near-lifesize pictures. This may have made it more likely that participants would think of these features differently, and perhaps as more two-dimensional than three-dimensional. However, to some extent this must be people's perception of geographic-scale objects anyway, even in cultures where maps are less prevalent, since they are inevitably much wider and longer than they are tall.

Where does this leave the hypothesis of scale-invariant spatial relations? Naturally, we should express caution about the generalisability of this study to more realistic contexts of use: like all experiments, there was some potential for artefactual results. However, it is hard to believe that an artificial context, and one in which the same illustration was used across scales (which should have encouraged participants to respond similarly as well), would show *more* scale variance than in a richer and more complex context. Further research is clearly needed to confirm and expand these findings: while it seems likely that some prepositions may maintain scale invariance, others do not. Prepositions such as *near* seem genuinely indifferent to scale precisely because they have no semantic specification of object properties (e.g., *line* *near a circle / planet near the sun*). However, we cannot conclude from such isolated examples that all such spatial language will be scale invariant.

Conversely, given the different physical entities and structures implied by the above analysis, it would be possible to maintain that scale invariance has not been disproved for situations where objects at different scales are more closely matched in their physical properties. However, our frustrating search for such equivalent examples across different scales and domains demonstrates in itself that the statement of scale invariance is ultimately meaningless. The physical world as experienced by humans, and described by human language, is not a fractal: scale changes its very physical nature, and hence the meaning of its spatial relations.

If this is the case, then it is arguably not scale *per se* that determines spatial preposition choices, but the nature and relative importance of the 'figure' and 'ground' objects which vary unavoidably with it. 'Ground' usually refers to the larger and more immobile object in a spatial relation, and to the object of a sentence or prepositional phrase, whereas the smaller object (and the sentence's subject) is more often the 'figure' [21]. The figure and ground objects of different scales will have different granularity, or degree of topological and metric precision. Some spatial prepositions also carry information that implies the objects' shape (e.g. line or area), such as *along* or *across*. Golledge [22] states that in describing topological relations, "the choice of nouns and prepositions conveys conventional information that is often fuzzy or inaccurate" (p.411). Yet convention has taught us which prepositions to use for certain concepts, which makes it possible for humans to understand spatial relations and spatial concepts. The present experiment demonstrates significant variance among spatial prepositions and their implied relations, suggesting that spatial prepositions strongly reflect people's reasoning about figure-ground relations in space.

Landau [21] has demonstrated that the characteristics of the 'ground' object tend to have a greater impact on spatial language use than those of the 'figure'. This appears to be borne out to some extent by the current study - it was apparently the nature of the <u>leaf</u>, as opposed to the nature of a <u>park</u> or <u>country</u>, that caused a problem for *through*. Although a <u>string</u> is as much a single three-dimensional solid object as a <u>leaf</u> is, whereas an <u>old road</u> is arguably more two-dimensional and 'pathlike', even a closer equivalent to a road (e.g. a slug trail) would not be deemed to run 'through' a leaf in the way that a road can run through a park. A vein within the leaf could do so, but this forms part of the leaf itself. Even if a road could be argued to be part of a park in a similar way, the fact that the traversal of a non-component object such as a car or highway can still be described with either *across* or *through* indicates that the difference is real. Similarly with the <u>gas pipeline</u> and the <u>country</u> (as opposed to the <u>string</u> and the <u>leaf</u> or the <u>old road</u> and the <u>park</u>), it was the nature of countries that seems to have caused a problem: no line object is expected to stop dead at a country's abstractly defined borders, whereas two *physically* distinct objects may well lie separately in space.

It should be noted that the finding that the same spatial terms are considered differently within English at different scales does not imply that this would generalise to all other languages. We know, for instance, that speakers of some languages apply to smaller scales some linguistic terms that in English we mostly[4] reserve for larger ones (e.g. [23]), although it is currently unclear whether this also applies to preposition use. However, the number of spatial prepositions in other languages varies greatly compared to English (e.g. [21]). Therefore we certainly cannot assume that scale is equally variant in every language: we might expect it to vary more where there is an atypically large and finely differentiated range of prepositions, as is the case for English and for other Germanic languages. Nevertheless, it seems likely (because of the physical difference in existent object types at different scales, as discussed) that some scale effects will still be found even in languages with reduced prepositional inventories. Alternatively, other grammatical and lexical markings which express distinctions similar to Germanic prepositions might be expected to exhibit scale variance.

Overall, although obviously further research is needed to extend these findings, this experiment has given us some insight into the variable status of some key spatial prepositions at different scales, and into some of the factors that may influence their use. The results lead to the conclusion that scale is not neutral as Talmy [13,14] suggested, but rather plays a significant role in linguistic topology. Reasons for these

[4] Since US and British English can show differences in usage of prepositions (e.g. Clark 1968), we also checked these results with a second sample of 11 US English speakers. We found the same pattern of main effects, although without any strong interactions; however, a combined ANOVA still suggested significance for the interactions as before, with no significant influence of English dialect. Overall this strengthens our view that scale and type of relation do affect people's expectations of spatial preposition use, independently of potential dialect variations.

differences may include humans' spatial experience, the abstractions that they use when reasoning about space, and the variable nature of our classification of the world into 'objects' at different scales.

References

1. Egenhofer M.J., Franzosa R.: Point-Set Topological Spatial Relations. International Journal of Geographical Information Systems 5 (1991) 161-174
2. Egenhofer M.J., Herring J. R.: Categorizing Binary Topological Relations Between Regions, Lines, and Points in Geographic Databases. Technical Report, Department of Surveying Engineering, University of Maine, Orono, Maine (1992)
3. Peuquet, D.J.: Representations of Geographic Space: toward a conceptual synthesis. Annals of the Association of American Geographers 78 (1988) 375-394
4. Hernandez D.: Relative Representation of Spatial Knowledge – The 2D Case. In: Mark D. M., Frank A. U. (eds.): Cognitive and Linguistic Aspects of Geographic Space. Kluwer, Dordrecht (1991) 373-385
5. Lynch, K.: The Image of the City. MIT Press, Cambridge MA, USA (1960)
6. Egenhofer, M.J., Mark, D.M.: Naïve Geography. In: Frank, A., Kuhn, W. (eds.): Spatial Information Theory – A Theoretical Basis for GIS. Lecture Notes in Computer Science, Vol. 988. Springer, Heidelberg-Berlin (1995) 1-15
7. Lautenschütz, A.-K.: Spatial Relations: the influence of scale, context and spatial relation in linguistic topology. Unpublished Master's thesis, Institut für Geographie, Westfälische Wilhelms-Universität Münster, Germany (2006)
8. Riedemann C.: Naming Topological Operators at GIS User Interfaces. In: Toppen F., Painho M. (eds.): 8th Conference on Geographic Information Science, Estoril, Portugal (2005) 307-315
9. Guarino, N.: Formal Ontology and Information Systems. In: N. Guarino (ed.): Formal Ontology in Information Systems. IOS Press, Trento, Italy (1998) 3-15
10. Schwering, A., Raubal, M.: Spatial Relations for Semantic Similarity Measurement. In: Akoka, J. et al. (eds.): Workshop Papers (2nd International Workshop on Conceptual Modeling for Geographic Information Systems, CoMoGIS 2005). Lecture Notes in Computer Science, Vol. 3770. Springer-Verlag, Berlin (2005) 259-269
11. Mark, D. M., Egenhofer, M. J.: Calibrating the Meanings of Spatial Predicates from Natural Language: Line-Region Relations. Sixth International Symposium on Spatial Data Handling, Edinburgh, Scotland (1994) 538-553
12. Coventry, K. R., Garrod, S. C.: Saying, Seeing and Acting: The Psychological Semantics of Spatial Prepositions. Essays in Cognitive Psychology Series. Psychology Press, Hove and New York (2004)
13. Talmy, L.: How language structures space. In: Pick, H.L., Acredolo, L.P. (eds.): Spatial Orientation – Theory, Research and Application. Plenum Press, New York (1983) 225-282
14. Talmy, L.: Toward a Cognitive Semantics. Volume I: Concept Structuring Systems. MIT Press, Cambridge, MA, USA (2000)
15. Mark, D.M., Frank, A.U., Egenhofer M.J., Freundschuh, S.M., McGranaghan, M., White, R.M.: Languages of Spatial Relations. Initiative Two specialist meeting report (NCGIA Technical Paper 89-2). National Center for Geographic Information and Analysis (1989)
16. Pederson, E.: How Many Reference Frames? In: Spatial Cognition III. Lecture Notes in Computer Science, Vol. 2865. Springer, Berlin (2003) 287-304
17. Montello, D.: Scale and Multiple Psychologies of Space. In: Frank, A.U., Campari, I. (eds.): Spatial Information Theory: Theoretical Basis for GIS. Lecture Notes in Computer Science, Vol. 716. Springer Verlag, Heidelberg-Berlin. (1993) 312-321
18. Hart, G., Temple, S., Mizen, H.: Tales of the River Bank: first thoughts in the development of a topographic ontology. In: Toppen, F., Prastacos, P. (eds.): Proceedings of 7th AGILE Conference. Crete University Press, Heraklion, Greece (2004) 165-168
19. Brugman, C.: The Story of "Over". Unpublished Master's thesis, Linguistics, University of California (1981)
20. Herskovits, A.: Language and Spatial Cognition: an interdisciplinary study of the prepositions in English. Cambridge University Press, Cambridge, UK (1986)
21. Landau, B.: Multiple Geometric Representations of Objects in Languages and Language Learners. In: Bloom, P., Peterson, M., Nadel, L., Garrett, M. (eds.): Language and Space. MIT Press, Cambridge MA, USA (1996) 317-63
22. Golledge R.G., Stimson R.J.: Spatial Behavior – A Geographic Perspective. The Guilford Press, New York (1997)
23. Pederson, E.: Geographic and Manipulable Space in Two Tamil Linguistic Systems. In: Frank, A.U., Campari, I. (eds.): Spatial Information Theory. Springer-Verlag, Berlin (1993) 294-311